ESD-TR-90-188

AD-A235 404

Technical Report
914

Binary Optics Technology: Theoretical Limits on the Diffraction Efficiency of Multilevel Diffractive Optical Elements

G.J. Swanson

1 March 1991

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

Prepared for the Defense Advanced Research Projects Agency
under Air Force Contract F19628-90-C-0002.

Approved for public release; distribution is unlimited.

This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Defense Advanced Research Projects Agency under Air Force Contract F19628-90-C-0002 (ARPA Order Number 5328).

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Hugh L. Southall

Hugh L. Southall, Lt. Col., USAF
Chief, ESD Lincoln Laboratory Project Office

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

BINARY OPTICS TECHNOLOGY: THEORETICAL LIMITS ON THE DIFFRACTION EFFICIENCY OF MULTILEVEL DIFFRACTIVE OPTICAL ELEMENTS

G.J. SWANSON
Group 52

TECHNICAL REPORT 914

1 MARCH 1991

LEXINGTON MASSACHUSETTS

ABSTRACT

Theoretical constraints limit the diffraction efficiency obtainable from multi-level diffractive optical elements. The scalar theory, commonly used to predict diffraction efficiencies, is overly optimistic. An extension to this theory is presented and compared with rigorous electromagnetic theory calculations. The extended scalar theory adds a degree of intuition in understanding why the diffraction efficiency of these elements is limited.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1. INTRODUCTION

Diffractive optical elements are being considered as potential solutions to a number of optical design problems that are difficult or impossible to solve with conventional refractive and reflective elements. Two unique characteristics of diffractive elements can be exploited; the first is the dispersion property. Diffractive structures bend light rays of longer wavelengths more than those of shorter wavelengths, which is the reverse of refractive materials; therefore, diffractive structures minimize or eliminate the dispersive effects of refractive materials.

The second unique characteristic is the relative ease with which arbitrary phase profiles can be implemented. Advances in both diamond turning technology and the use of semiconductor fabrication equipment have made possible the construction of a variety of diffractive elements. Diamond turning technology allows fabricating diffractive surfaces over large areas in a relatively short period of time. However, there are limitations: the phase profile has to be circularly symmetric, and the accuracy with which a diffractive profile can be made is dependent on the tip size of the diamond turning tool.

Using semiconductor fabrication equipment to make diffractive elements has become a powerful technique. This particular approach produces a stepped approximation, referred to as a "multilevel structure," to the ideal profile. As the number of levels becomes large, the diffractive structure approaches the continuous profile. Diffractive elements can be made with feature sizes down to 0.5 μm. The diffractive profiles can be very general with no symmetry restrictions, for example, lenslet arrays, which are being used to increase the collection efficiency of detector arrays and as components of wavefront sensing devices. These arrays are composed of individual diffractive lens profiles that are corrected for spherical aberration. Each lens has a rectangular aperture so that 100% of the area is covered. Such lenslet arrays would be difficult to fabricate any other way.

The diffractive optical elements that are fabricated by diamond turning or by using semiconductor fabrication equipment are surface relief elements. Surface relief diffractive elements are a particular class of diffractive elements that impart a phase delay to an incident wavefront in a very thin layer close to the surface of the element. The thickness of this layer is on the order of the incident wavelength. The phase delay is imparted to the incident wavefront by selectively removing material from the surface of the substrate.

Diffractive optical elements are different from reflective or refractive elements in that a light ray incident on a diffractive element is split into many rays, only one of which travels in the desired direction; its magnitude, relative to the sum of the magnitudes of all the split light rays, is called the diffraction efficiency. In most cases, a diffraction efficiency of one is desired, which is equivalent to all the light traveling in the chosen direction.

The diffraction efficiency that can be expected in practice from a particular diffractive element is limited by theory as well as by fabrication tolerances. The ability to fabricate diffractive elements has improved dramatically over the past few years — so much so that the attainable diffraction

efficiency for many elements (particularly those operating in the far infrared) is limited almost exclusively by theory. Performance degradation of diffractive optical elements due to fabrication errors has been investigated by others [1,2]. This report concentrates on the strictly theoretical limitations of achievable diffraction efficiency. It is, therefore, assumed that the surface relief profiles can be fabricated with infinite accuracy. The resulting diffraction efficiency calculations place a theoretical upper limit on attainable performance.

Whether a diffractive element will work for a particular application is ultimately determined by the obtainable diffraction efficiency; for example, consider the case of a lenslet array that is used to increase the light-gathering ability of a detector array. Certain detector arrays are made with a substantial fraction of dead space on the detector plane. A lens, properly placed in front of each detector, would effectively concentrate the light that would have fallen on the dead space onto the detector. For typical detector arrays under consideration, the increase in light-gathering capacity that a lenslet array can achieve is about a factor of 4, assuming that the lenslets have a diffraction efficiency of 100%. If the diffraction efficiency were only 50%, the increase in light-gathering efficiency would be only a factor of 2. If the diffraction efficiency dropped to 25%, the lenslet array would contribute absolutely nothing. Therefore, the diffraction efficiency that can reasonably be expected from a diffractive element is an important parameter.

Conventional lens design programs are now commonly used to model and optimize diffractive phase profiles. These lens design codes assume that the diffraction efficiency of a diffractive element is 100%. These codes are capable of determining phase profiles, but obtainable diffraction efficiency has to be determined separately. Theoretically, diffraction efficiency is a function of a number of parameters: the index of refraction of the substrate, the size of the zones of the diffractive profile relative to the incident wavelength, the polarization and angle of incidence of the incident light, and the depth and shape of the surface profile within a zone.

In theory, Maxwell's equations can determine exactly the diffraction efficiency of any diffractive structure. In practice, it is not possible to obtain exact solutions for the majority of cases. Numerical solutions are possible for certain diffractive structures; however, the necessary algorithms are very computationally intensive.

One of the simplest and most widely used ways to predict diffraction efficiencies is to use a scalar theory. The scalar theory of diffraction from a surface relief structure is based on a simplification of Maxwell's equations and a simplified model of the surface relief structure. The region of validity of the scalar theory is in the limit of the wavelength-to-zone spacing approaching zero. In other words, the size of the diffracting feature has to be very large compared with a wavelength of the incident light. The light is, therefore, deviated from the incident direction by a small angle. Section 2 describes the scalar theory and uses it to predict diffraction efficiency.

When the ratio of the wavelength-to-zone spacing approaches one, the incident light is deviated by large angles approaching 90 deg. It is in this regime that the scalar theory completely breaks down. Reliable estimates of diffraction efficiency can no longer be obtained from the scalar theory; however, numerical solutions to Maxwell's equations can be obtained for periodic diffracting

structures, i.e., gratings. If the grating period becomes much larger than a few wavelengths, the algorithm becomes too computationally intensive. Section 3 describes briefly the electromagnetic theory approach used to solve Maxwell's equations numerically for periodic structures.

In determining the diffraction efficiency of a grating, the scalar theory is valid for large period-to-wavelength ratios while the electromagnetic theory can only be used for very small period-to-wavelength ratios. A large void is left between the two limits where the scalar theory is not very accurate and the electromagnetic theory is numerically prohibitive. An approach to obtaining more reliable results for the diffraction efficiency in this region of period-to-wavelength ratios is to extend the scalar theory. This extended theory, developed in Section 4, combines aspects of geometrical optics with conventional scalar theory.

Section 5 compares the results of the three theories for a few representative examples, and the consequences of the theoretically obtainable diffraction efficiency for various applications are discussed.

2. SCALAR THEORY OF DIFFRACTION EFFICIENCY

The scalar theory of diffraction is based on the assumptions that light can be treated as a scalar rather than vector field and that the electric and magnetic field components are uncoupled. Two conditions are commonly stated as necessary for the scalar theory to have any validity: the size of the diffracting features must be large compared to the incident wavelength, and the diffracted field must be observed far from the diffracting structures [3].

A further approximation, referred to as the "Fresnel approximation," allows an integral solution of the propagation of the light field. The Fresnel approximation assumes that spherical waves can be approximated by quadratic waves. Within the realm of Fresnel diffraction, given the light field at some initial plane, the light field can be determined at any plane. Mathematically, the process of Fresnel diffraction is expressed by

$$U(x,y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy U(x_0, y_0) \exp\{\frac{i\pi}{\lambda z}[(x-x_0)^2 + (y-y_0)^2]\}, \tag{1}$$

where the initial light field, $U(x_0, y_0)$, is propagated a distance z, resulting in the light field $U(x,y)$. Multiplicative factors preceding the integral are generally not important and are omitted.

If the propagation distance is large enough so that the quadratic phase term in the integral of Equation (1) can be ignored, the resulting expression, again neglecting the unimportant multiplicative factors, becomes

$$U(f_x, f_y) = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy U(x_0, y_0) \exp\{-i2\pi[f_x x_0 + f_y y_0]\}, \tag{2}$$

where $f_x = x/\lambda z$ and $f_y = y/\lambda z$. Equation (2) represents the approximation known as the Fraunhofer diffraction and is the foundation for calculating diffraction efficiencies of surface relief diffractive elements in the scalar regime. For a periodic structure, i.e., grating, the amplitudes of the various diffraction orders can be determined by a simple Fourier transformation of the grating transmittance function. This simplification will be used to calculate the theoretical performance of multilevel phase gratings. It should be noted that in the scalar theory, the diffraction efficiency of an arbitrary diffractive optical element can be directly related to the diffraction efficiency of a grating [4]. It is, therefore, only necessary to determine the diffraction efficiency of a grating structure.

2.1 Diffraction Efficiency of a Multilevel Phase Grating

The surface relief profile of a one-dimensional, multilevel phase grating is shown in Figure 1. In order to calculate the diffraction efficiency of this grating structure, the far-field of one grating period has to be determined. The transmittance function of one period can be described by the

summation of the transmittances of N subperiods of width T/N, where N is the number of phase levels within one period of dimension T.

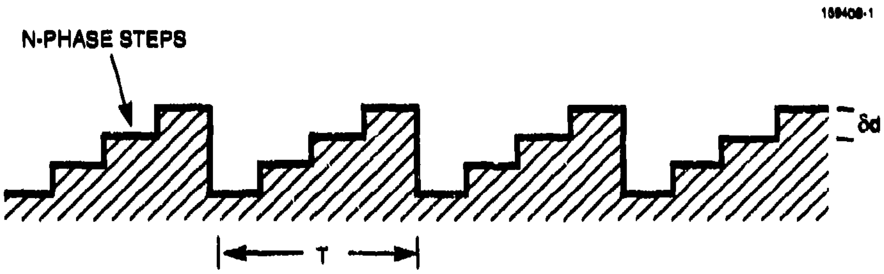


Figure 1. Surface relief profile of a one-dimensional, multilevel phase grating.

Each subperiod is a rect function of width T/N, centered at $x = (m + 1/2)T/N$, where m is an integer from 0 to $N - 1$. In the scalar approximation, the phase delay imparted by each subperiod can be expressed as $\phi = m\phi_0/N$, where ϕ_0 is the largest phase delay of all subperiods.

The far-field amplitude distribution of a subperiod, centered at a position x, with a width T/N and a phase delay of ϕ, can be calculated from Equation (2) with the result

$$U'(f) = \frac{\sin(\pi Tf/N)}{\pi Tf/N} \exp\{-i2\pi xf\} \exp\{i2\pi\phi\}. \tag{3}$$

The far-field amplitude distribution of a total period can then be expressed as a summation of the far-field amplitude distributions of the N subperiods within the total period:

$$U(f) = \frac{1}{N} \sum_{m=0}^{N-1} \frac{\sin(\pi Tf/N)}{\pi Tf/N} \exp\{-i2\pi((m + \frac{1}{2})T/N)f\} \exp\{i2\pi m\phi_0/N\}. \tag{4}$$

Repeating the period an infinite number of times constrains the far-field to have nonzero values only at positions $f = l/T$, where l is an integer that represents the lth diffraction order. The far-field amplitude of the lth diffraction order can be written as

$$A_l = \exp\{-i\pi l/N\} \frac{\sin(\pi l/N)}{\pi l/N} (1/N) \sum_{m=0}^{N-1} \exp\{-i2\pi(l - \phi_0)m/N\}. \tag{5}$$

The diffraction efficiency η_l of the lth order is $A_l A_l^*$,

$$\eta_l = \frac{\sin^2(\pi l/N)}{(\pi l/N)^2} 1/N^2 [\sum_{m=0}^{N-1} \exp\{-i2\pi(l-\phi_0)m/N\}]^2. \tag{6}$$

The summation in Equation (6) can be readily evaluated

$$[\sum_{m=0}^{N-1} \exp\{-i2\pi(l-\phi_0)m/N\}]^2 = \frac{\sin^2(\pi(l-\phi_0))}{\sin^2(\pi(l-\phi_0)/N)}. \tag{7}$$

Substituting the result of Equation (7) into (6) gives the expression for the diffraction efficiency of the lth order as

$$\eta_l^N = [\frac{\sin(\pi(l-\phi_0))}{\pi l} \frac{\sin(\pi l/N)}{\sin(\pi(l-\phi_0)/N)}]^2, \tag{8}$$

where N is the number of phase levels, $\phi_0 = N\phi$, and ϕ is the phase depth change in waves of one subperiod.

Equation (8) is the basis for calculating diffraction efficiencies of surface relief diffractive optical elements. Within the scalar theory region of validity, this equation can determine the amount of light in any diffraction order for any number of phase levels. Equation (8) shows that for a given number of phase levels, N, the diffraction efficiency of the lth diffraction order is a function of one parameter, ϕ_0. This ϕ_0 parameter can be related to the physical step height of a multilevel structure, as well as the incident wavelength and the angle of incidence of light impinging on the diffractive surface.

Figure 2 illustrates the relationship between the parameters necessary to define ϕ_0 in terms of physical properties. Two light rays are shown impinging on two neighboring subperiods in a multilevel structure. The index of refraction of the diffractive element is n and the angle of incidence is θ_1. The physical step height between the neighboring subperiods is δd.

The parameter ϕ, previously defined as the phase difference in waves between two neighboring subperiods, is therefore defined in terms of the parameters of Figure 2 as

$$\phi = \frac{1}{\lambda}(ny_1 - y_2), \tag{9}$$

where the distances y_1 and y_2 are geometrically determined to be

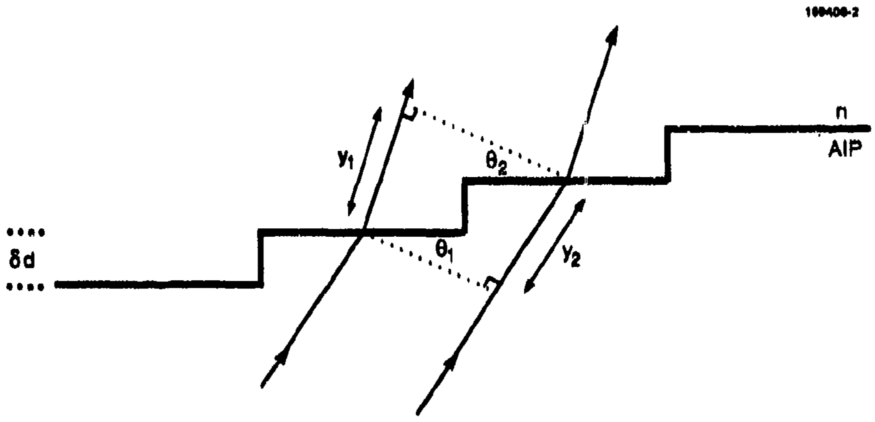


Figure 2. Light rays traced through two neighboring subperiods.

$$y_1 = \frac{\delta d}{\cos\theta_2} + \sin\theta_2(x - \delta d\tan\theta_2) \tag{10}$$

and

$$y_2 = \frac{\delta d}{\cos\theta_1} + \sin\theta_1(x - \delta d\tan\theta_1). \tag{11}$$

Inserting Equations (10) and (11) into (9) results in an expression for ϕ that can be trigonometrically reduced to

$$\phi = \frac{\delta d}{\lambda}\{n\cos\theta_2 - \cos\theta_1\}. \tag{12}$$

Relating θ_1 to θ_2 in Equation (12) through Snell's law results in the following expression for ϕ as a function of the step height, the index of refraction of the substrate, and the angle of incidence in air:

$$\phi = \frac{\delta d}{\lambda}[\sqrt{n^2 - \sin^2\theta_1} - \cos\theta_1]. \tag{13}$$

The parameter ϕ_0 in Equation (8) is, again, $\phi_0 = N\phi$. It should also be noted that for the case of normal incidence θ_1 becomes zero, and Equation (13) reduces to the simple expression

$$\phi = \delta d(n-1)/\lambda. \tag{14}$$

2.1.1 Examples

Equation (8) is a general scalar theory expression used to determine the diffraction efficiency of multilevel diffractive elements. An equivalent scalar theory expression for continuous profile diffractive elements, such as those fabricated by diamond turning techniques, can be found by taking the limit of Equation (8) as the number of levels N approaches infinity. The resulting expression for an infinite number of phase levels becomes

$$\eta_l^{\infty} = [\frac{\sin(\pi(l-\phi_0))}{(\pi(l-\phi_0))}]^2. \tag{15}$$

Notice that the diffraction efficiency of the first order, η_1^{∞}, becomes 100% when $\phi_0 = 1$. This is the result of the scalar theory that claims that 100% diffraction efficiency is possible.

The first diffraction order is usually of most interest and usually requires the highest diffraction efficiency. The diffraction efficiency of the first order is maximum when $\phi_0 = 1$. The first-order diffraction efficiency of an optimized N-level element can be found by setting l and ϕ_0 both equal to one:

$$\eta_1^N = [\frac{\sin(\pi/N)}{(\pi/N)}, \tag{16}$$

expressing the maximum first-order diffraction efficiency one can expect from an N-level element in the scalar approximation.

The ϕ_0 parameter can be expressed as a function of the total depth d of the diffractive profile rather than the depth δd of a subperiod. The total depth d is simply related to δd, by $d = (N-1)\delta d$. The ϕ_0 parameter, for normal illumination, becomes

$$\phi_0 = (\frac{N}{N-1})\frac{(n-1)}{\lambda}d. \tag{17}$$

Setting ϕ_0 equal to one determines the optimum total depth for an N-level diffractive profile on a substrate of index n, to be used at a wavelength λ:

$$d = \frac{(N-1)}{N}\frac{\lambda}{(n-1)}. \tag{18}$$

In the limit of the number of levels approaching infinity, the well-known expression for the optimum depth, $d = \lambda/(n-1)$, is obtained.

It is a fact that the diffraction efficiency of a diffractive structure is wavelength dependent. From the previous analysis, it can be deduced that the optimum step height for normal incidence and wavelength λ_0 is

$$\delta d = \frac{\lambda_0}{N(n-1)}. \tag{19}$$

Substituting Equation (19) into (13) results in an expression for ϕ from which ϕ_0 can be determined to be

$$\phi_0 = \frac{\lambda_0}{\lambda}\left[\frac{\sqrt{n^2 - \sin^2\theta_1} - \cos\theta_1}{(n-1)}\right]. \tag{20}$$

Equation (8), in conjunction with (20), can be used to determine the diffraction efficiency of an N-level element as a function of wavelength and incident angle, for which the first-order diffraction efficiency has been maximized for wavelength λ_0 and normal incidence.

Figure 3 plots the first-order diffraction efficiency as a function of wavelength for various values of N. The element was optimized, as described above, to have a maximum diffraction efficiency at wavelength λ_0 and normal incidence.

Figure 4 plots the first-order diffraction efficiency at wavelength λ_0 as a function of incident angle for various values of N. The element was optimized to have a maximum diffraction efficiency at wavelength λ_0 and normal incidence. The figure reveals that in the scalar approximation the diffraction efficiency of these elements is very insensitive to the angle of incidence. This result reflects positively on the concept of placing diffractive surfaces on refractive optical elements, with the intent that the diffractive surface minimizes the aberrations of the refractive element. In such cases, the period-to-wavelength ratio of the diffractive structure is usually large, lending credibility to the scalar approximations; however, the range of incident angles impinging on the diffractive surface becomes quite large. Figure 4 shows that the diffraction efficiency, in general, will not suffer very much as a consequence of the large range of input angles.

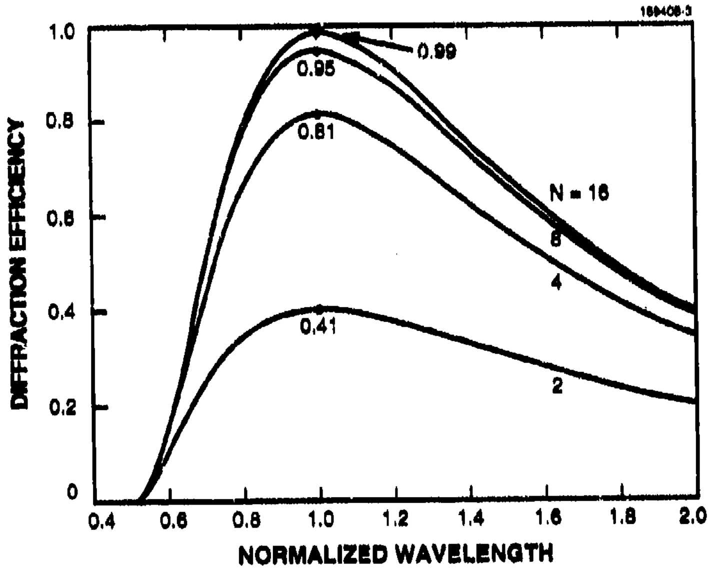


Figure 3. The first-order diffraction efficiency as a function of wavelength and the number of phase levels.

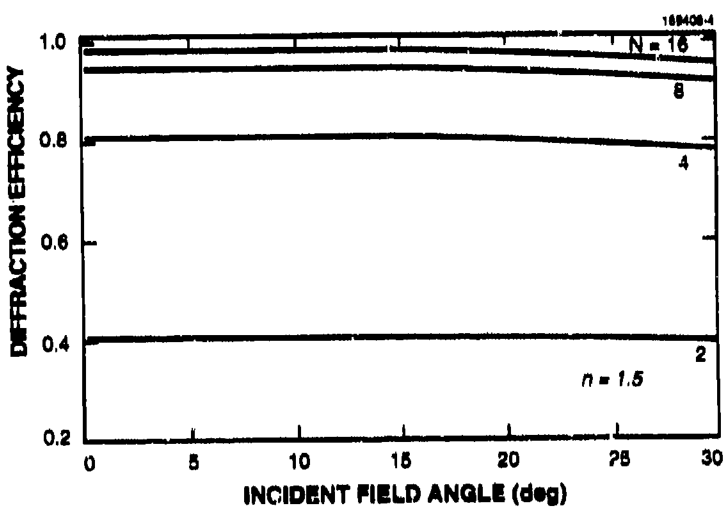


Figure 4. The first-order diffraction efficiency as a function of incident angle.

3. RIGOROUS ELECTROMAGNETIC THEORY OF DIFFRACTION EFFICIENCY

In Section 2, analytical expressions for the diffraction efficiency of surface relief phase gratings were developed using a scalar theory. As mentioned earlier, the diffraction efficiency of structures more complex than simple gratings can be directly related to the diffraction efficiency of the gratings through a nonlinear limiter analysis [4]. This allows a closed-form solution of the diffraction efficiency for any surface relief diffractive optical element.

The scalar theory is useful for designing surface relief diffractive elements with periods that are much larger than the wavelength for which the element is to be used. When the periods on the diffractive element become comparable in magnitude to the wavelength, the scalar theory (developed in Section 2) gives unreliable values for diffraction efficiency. The amount of discrepancy between the diffraction efficiency predictions of the scalar theory and reality is a function of the period-to-wavelength ratio and the index of refraction of the substrate.

In order to get a more reliable prediction of expected diffraction efficiencies, a more accurate theory must be used. In principle, Maxwell's equations could be solved for a particular diffractive structure, giving results that would be extremely accurate. In practice, the solutions to Maxwell's equations have to be calculated numerically.

Various approaches to solving the electromagnetic equations of grating diffraction exist. Although they are equally valid, this report uses the approach first employed by Moharam and Gaylord [5], which is based on a coupled wave theory approach to solving Maxwell's equations. A brief outline follows. (Because the details are too numerous to discuss in this report, the reader is referred to Reference 5.)

An electromagnetic field incident on a phase grating can be divided into three main regions. The first, described by a homogeneous permittivity ϵ_1, is where the incident and reflected fields propagate. The second is the modulation region of the grating profile, with permittivity alternating between ϵ_1 and ϵ_3, the permittivity of the third region. This third region is where the transmitted field propagates and is characterized by the homogeneous permittivity ϵ_3. In all three regions, permeability is equal to the permeability of free space.

The electromagnetic fields in the first and third regions can be expanded as sums of plane waves with the wave vectors determined from the Floquet condition. In the second region, the electromagnetic fields are expressed as Fourier expansions of the space harmonic fields. The second region is divided into N layers of equal thickness, each represented by the characteristics of the grating at the middle of the layer. The permittivity of each layer can be represented by a Fourier expansion. The permittivity in the second region, ϵ_2, alternates within a layer between ϵ_1 and ϵ_3.

The solution for the amplitudes of the reflected and transmitted diffraction orders is achieved by applying Maxwell's equations at the boundaries between the N layers. The electric and magnetic fields must have continuous tangential components.

An extensive computer code, DIFFRACT, has been developed based on the coupled wave theory. The accuracy of the code is dependent on the number of layers used to describe the grating modulation region and the number of orders retained in the Fourier expansion of the electromagnetic fields. The computation time necessary to solve for the diffraction efficiency increases linearly with the number of layers. In other words, the amount of computer time used to solve an N layer grating structure is twice that of an $N/2$.

The computation time necessary to solve for a grating is proportional to the cube of the number of orders retained in the Fourier expansion. In order to obtain an accurate solution, all the propagating orders, as well as a few evanescent orders, should be retained. The number of propagating orders from a grating is determined by the period-to-wavelength ratio; the larger the ratio, the more propagating diffraction orders. The computation time is, therefore, a strong function of the period-to-wavelength ratio. Furthermore, the maximum period-to-wavelength ratio grating that can reasonably be solved is dependent on the available computing power. In general, gratings with period-to-wavelength ratios greater than 10 become unreasonable to try to solve using this algorithm.

As seen above, one of the main constraints of the rigorous coupled wave theory, as well as other rigorous electromagnetic theories, is the limit on the maximum period-to-wavelength ratio grating that can be solved. The scalar theory, on the other hand, is only valid in the very large period-to-wavelength regime. A void remains between the usefulness of the two theories where unfortunately, a large percentage of the diffractive structures are being considered for various applications.

Another property of the rigorous electromagnetic theory is that it lends itself to very little intuitive insight into what to expect for diffraction efficiencies from gratings. Section 4 presents an intermediate theory for multilevel diffractive optical elements that attempts to bridge the gap between the scalar and the rigorous electromagnetic theories. This intermediate theory partially explains, in an intuitive fashion, the falloff of diffraction efficiency as a function of period-to-wavelength ratio.

4. EXTENDED SCALAR THEORY OF DIFFRACTION EFFICIENCY

The scalar theory of diffraction, as described in Section 3, is valid only for diffractive structures that have very large period-to-wavelength ratios. The rigorous electromagnetic theories of grating diffraction allow numerical solutions for only small period-to-wavelength ratios due to the computational complexity of the algorithms. A useful theory would function in the region of intermediate values of period-to-wavelength ratios, would be more accurate than the scalar theory, and would be computationally simpler than the rigorous electromagnetic theories.

The intermediate theory presented here, called the extended scalar theory, is like the scalar because it is strictly valid only in the confines of very large period-to-wavelength ratios, but for intermediate values of period to wavelength, agreement with reality is much better.

The major assumption that the extended scalar theory attempts to avoid is that the phase delay of the incident light, caused by the grating, occurs in an infinitely thin layer. The effects of the finite thickness of the grating profile are taken into consideration.

The finite thickness of the grating profile is treated by combining the scalar theory (based on wave propagation) with a geometrical theory (based on ray tracing). The incident light field is assumed to propagate through the thickness of the grating profile according to geometrical optics. Once the light exits the grating profile, the scalar theory based on wave propagation is applied.

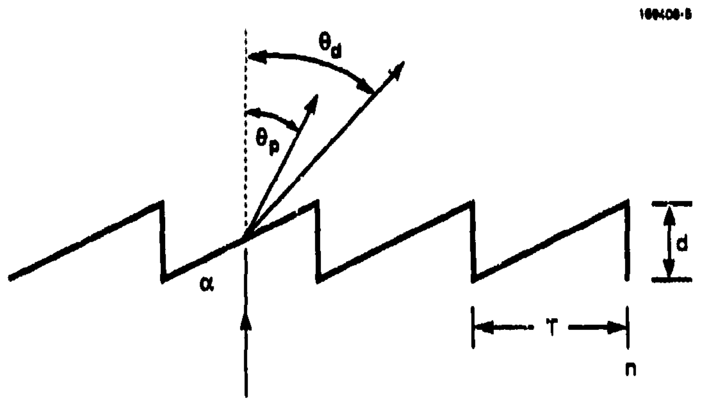
4.1 Optimum Grating Profile Depth

As mentioned above, the most widely used scalar theory assumes that the phase delay associated with a surface relief phase grating occurs in an infinitely thin layer on the surface of the substrate. This phase delay is physically implemented, however, by etching away certain areas of the substrate surface. The phase delay is the result of the optical path length difference due to the variation in surface profile thickness. The conversion of a phase delay into a physical thickness for a diffractive element designed to have a maximum first-order diffraction efficiency was shown in Section 2 to result in a physical depth of d, where $d = \lambda/(n-1)$. Notice that the optimum depth based on the scalar theory is only a function of the wavelength and index of refraction of the substrate.

The mathematical assumption that the phase delay occurs in an infinitely thin layer is obviously unrealistic. Only for the case of substrates with extremely large refractive indices would the theory begin to agree with reality. Therefore, the scalar value of depth d is also an approximation. The questions "How bad is the assumption of the scalar theory?" and "What is the actual optimum depth?" need to be answered.

The approach used to determine the optimum depth by extending the scalar theory is shown in Figure 5 for the case of light normally incident on the substrate boundary and traveling from the substrate into air. The angle, θ_d, at which the first diffraction order travels from the grating is simply determined by the grating equation

$$\sin \theta_d = \lambda / T. \tag{21}$$



- SNELLS LAW: $n \sin(\alpha) = \sin(\theta_p + \alpha)$
- GRATING EQUATION: $\sin \theta_d = \frac{\lambda}{T}$
- SET $\theta_d = \theta_p$
- SOLVE FOR d

Figure 5. Geometrical ray trace through a surface relief grating.

If one now considers each period of the grating to consist of a miniature refractive prism, light rays can be traced geometrically through each facet. The angle that the light rays exit the prism, θ_p, is simply governed by Snell's law

$$n \sin \alpha = \sin(\theta_p + \alpha), \tag{22}$$

where $\alpha = \arctan d/T$.

An intuitive argument would suggest that the first diffraction order will have its maximum efficiency when the angle of the light rays traced through the prism θ_p is equal to the angle of the first diffraction order θ_d. The result of setting θ_p equal to θ_d and solving for d is

$$d = \frac{\lambda}{n - \sqrt{1 - (\lambda/T)^2}}. \tag{23}$$

Notice that this value of the grating depth is different from the scalar theory value. The most apparent difference is that the optimum depth given in Equation (23) is a function of the grating period, whereas the scalar theory value is independent of it. This immediately implies that for structures more complicated than gratings, the depth of the diffractive profile should vary as a function of the local period of the structure. Furthermore, it is worth noting that in the limit of the period T going to infinity, Equation (23) reduces to the scalar theory value.

From this point on, the depth value determined from Equation (23) is referred to as the "optimum depth" and represented by d_{opt}. The scalar depth value is represented by d_{app}. In order to see how the optimum varies from the scalar theory depth, it is useful to plot the ratio of the two as a function of the period-to-wavelength ratio, as shown in Figure 6 for two values of the index of refraction of the substrate. As expected, the ratio of d_{opt}/d_{app} asymptotically approaches a value of one as the period-to-wavelength ratio increases. The depth ratio deviates significantly from a value of one at small period-to-wavelength ratios. The exact period-to-wavelength ratio at which the deviation becomes significant is dependent on the index of refraction of the substrate. For high index of refraction substrates, the deviation occurs at smaller period-to-wavelength ratios than for low index of refraction substrates.

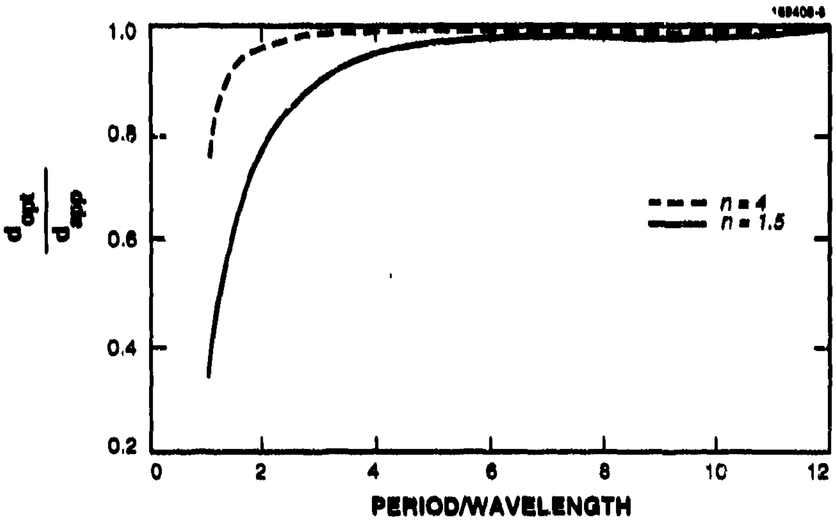


Figure 6. Extended scalar theory prediction of optimum depth as a function of the period-to-wavelength ratio.

Equation (23) was derived for normal incidence on the substrate boundary with the light traveling from the substrate into air. A more general expression for the optimum depth can be

derived using a similar approach to that used to derive Equation (23). Again, the idea is to simply equate the diffraction angle of the grating to the deviation angle of the prism for an arbitrary angle of incidence. The result of such an approach is the expression for the optimum depth as a function of incident angle as well as the wavelength-to-period ratio and the index of refraction:

$$d_{opt} = \frac{\lambda}{n\sqrt{1-(\sin\theta_i)^2} - \sqrt{1-(\frac{\lambda}{T} + n\sin\theta_i)^2}}. \tag{24}$$

Notice that Equation (24) reduces to (23) when the incident angle θ_i is set equal to zero. Equation (24) can also be used to determine the optimum depth for normal illumination when the light is traveling from air into the substrate. In Equation (24), θ_i is defined as the incident angle in the substrate material. For the case of normal illumination from air into the substrate, $\sin\theta_i$ has to be set equal to $-\frac{\lambda}{nT}$. The result is the optimum depth for normal incidence traveling from air into the substrate:

$$d_{opt} = \frac{\lambda}{n\sqrt{1-(\lambda/nT)^2} - 1}. \tag{25}$$

For all cases as the wavelength-to-period ratio approaches zero, the depth approaches the scalar theory value of $d_{app} = \lambda/(n-1)$.

The depth values determined above were based on a somewhat intuitive argument. There is no proof that the expressions derived determine the depth that results in a maximum first-order diffraction efficiency. To test these extended scalar theory depth values, the DIFFRACT program (described in Section 3) was used to calculate the theoretical first-order diffraction efficiency for various wavelength-to-period ratio gratings. The minimum ratio tested was 0.5, corresponding to a 30-deg diffraction angle for the first order. Calculations were done for both high- ($n = 4$) and low-index ($n = 1.5$) substrates. The depth of the gratings was varied over a region that included the optimum as well as the scalar theory depth. In all cases, the first-order diffraction efficiency was maximized when the depth was near that predicted by the extended scalar theory.

Figures 7 and 8 plot the first-order diffraction efficiency as a function of the wavelength-to-period ratio. Curves are plotted for gratings having depth values equal to both the scalar theory and the optimum. Figure 7 plots a substrate with a low index of refraction ($n = 1.5$), and Figure 8 plots a substrate with a high index ($n = 4$). The calculations for the high-index substrate include a single layer antireflection coating; the low-index substrate had none. In all cases, the optimum depth value, as predicted using the extended scalar theory, results in a higher diffraction efficiency than that predicted using the scalar theory.

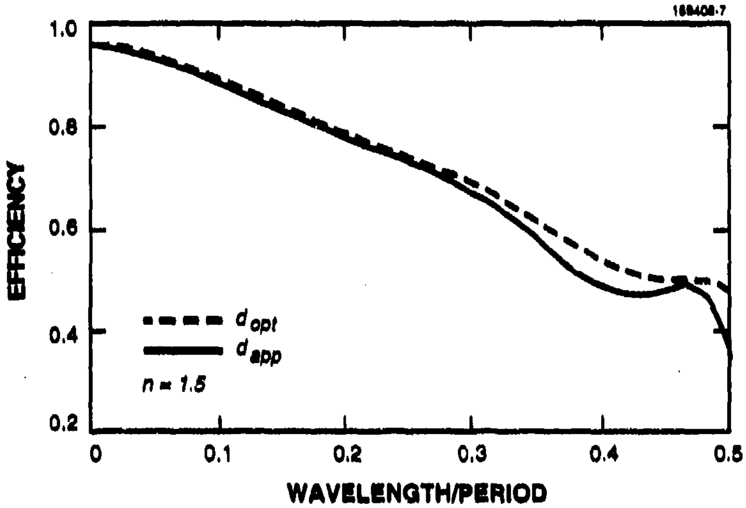


Figure 7. First-order diffraction efficiency as a function of the wavelength-to-period ratio for an $n = 1.5$ substrate.

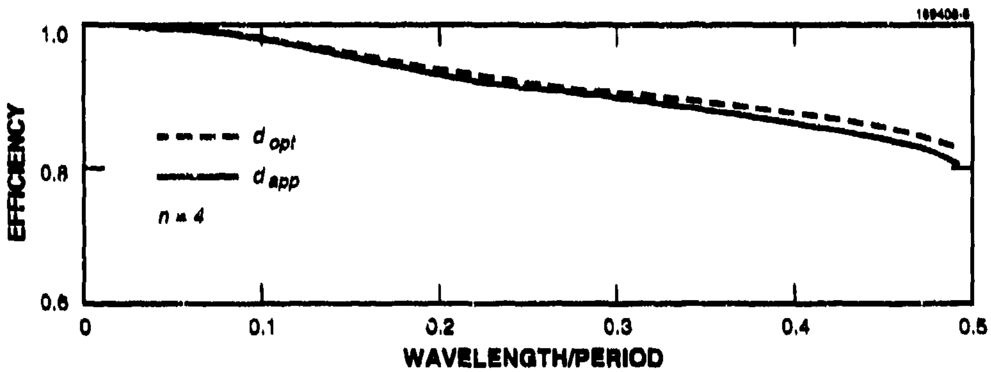


Figure 8. First-order diffraction efficiency as a function of the wavelength-to-period ratio for an $n = 4$ substrate.

4.2 Extending Scalar Theory Prediction of Diffraction Efficiency

Diffraction efficiency predictions based on the scalar theory are completely independent of the wavelength-to-period ratio. Figures 7 and 8 clearly show, however, that the diffraction efficiency is a function of the wavelength-to-period ratio. One of the major reasons that the scalar theory fails to predict this falloff is, again, largely due to the assumption that the phase delay occurs in an infinitely thin boundary of the substrate.

The concept of geometrically tracing rays through the finite depth of the diffractive structure and subsequently applying the scalar theory can be used to extend the prediction of diffraction efficiency. This approach, though obviously not an exact solution to the diffraction problem, is more consistent with the electromagnetic theory calculations.

The most apparent feature that emerges from geometrically tracing rays through the depth of the diffractive structure is an effect referred to as "light shadowing." Figure 9 illustrates the geometrical ray trace and shows the light shadowing resulting from a finite thickness structure. Light rays traveling in a direction normal to the substrate boundary are refracted at the substrate/air interface. The angle that the light rays deviate is determined from Snell's law. The depth d is assumed to be the value determined in Section 4.1 that optimizes the first-order diffraction efficiency. The period of the grating is T, and the index of refraction of the substrate is n.

The light rays that exit the grating structure in the first diffracted order no longer fill the entire grating area. Immediately after the grating, the ratio of the area filled with light to the total area is called the duty cycle (DC) and is equal to $\Delta T/T$. From a geometrical construction, the DC for the case illustrated in Figure 9 can be expressed as

$$\mathrm{DC} = 1 - \frac{d\lambda}{T^2\sqrt{1-(\lambda/T)^2}}. \tag{26}$$

Once the light rays are traced through the grating profile and the DC of the first diffraction order is determined, the scalar theory is applied to the exiting field. The light in the first diffraction order immediately after the grating resembles an unfilled aperture. It is a well-known result of the scalar theory that the amount of light that travels undiffracted through an unfilled aperture is equal to the DC of the unfilled aperture.

The light that is traced through one period of the grating encounters a stepped profile if the grating is made in a multilevel fabrication process. For this case, a fraction of the incident light equal to the DC given by Equation (26) is lost. Therefore, the fraction of light that resides in the first diffraction order can be approximately expressed by the product of the DC squared and the efficiency predicted from the scalar theory. Note from Equation (26) that the DC and, therefore, the first-order diffraction efficiency, is a function of the wavelength-to-period ratio; going to zero, the DC approaches one, and the first-order diffraction efficiency approaches the scalar theory value.

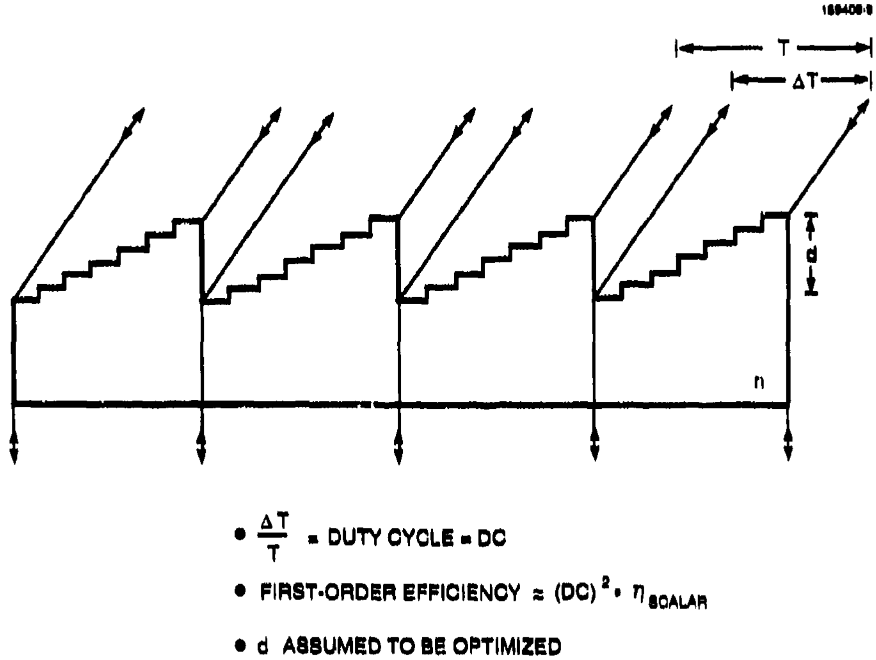


Figure 9. Light shadowing caused by finite depth surface relief profile.

A further extension could be approximated by including polarization effects. The scalar theory and its extension are polarization independent. These effects could be added to the extended scalar theory by including losses at the grating facet boundaries due to Fresnel reflection losses.

The extended theory is designed to be strictly valid only in the large period-to-wavelength ratio limit, as is the scalar theory, and more accurate for moderate wavelength-to-period ratios. As the period-to-wavelength ratio decreases, the extended scalar theory breaks down. The theory completely breaks down for a given index of refraction at the point where the slope of the individual facets within one period become large enough so that a light ray traced at the boundary will suffer from total internal reflection. Combining the equations for total internal reflection and the optimum grating depth results in an upper limit on the wavelength-to-period ratio for which extended scalar theory has any validity. This upper limit is expressed as

$$\left(\frac{\lambda}{T}\right)_{max} = \sqrt{1 - 1/n^2}; \tag{27}$$

for example, the extended scalar theory for a substrate with an index of refraction equal to 4 will totally break down when the wavelength-to-period ratio is equal t^ 0.97. For a substrate with a 1.5 index of refraction, the breakdown occurs at a wavelength-to-period ratio of 0.74. Section 5 compares the extended scalar theory with rigorous electromagnetic calculations. The maximum value of the wavelength-to-period ratio used in these comparisons is 0.5.

5. COMPARISON OF SCALAR, EXTENDED SCALAR, AND ELECTROMAGNETIC THEORIES

Three theories have been presented that can predict the diffraction efficiency from diffractive optical elements; each has strong points and weaknesses, and each complements the other in terms of information.

Obviously, the electromagnetic theory results in an exact solution to the problem of diffraction from a grating. Solutions to the electromagnetic theory can only be calculated numerically and computation time increases rapidly as the period-to-wavelength ratio increases; thus, there are two limitations. The first is the upper bound on the period-to-wavelength ratio for which a solution can be calculated, which is a function of the computer speed and how long one is willing to wait for the solution. The second limitation is the lack of any real insight into trying to optimize the diffraction efficiency of a diffractive structure.

The scalar theory is the least accurate yet easiest to use of the three; it allows for analytical expressions for the diffraction efficiency as a function of physical parameters. The analytical expressions give an insight into the design and/or feasibility of diffractive optical elements for a particular application. The diffraction efficiency calculated using the scalar theory is completely independent of the period-to-wavelength ratio. The value calculated can be used, however, as an upper bound on the obtainable diffraction efficiency. Scalar theory accuracy increases as the period-to-wavelength ratio increases. Thus, the theory becomes valid when the electromagnetic theory cannot be used due to computation time.

The extended scalar theory fills the void between the scalar and the electromagnetic. It retains the closed-form solution of the scalar theory and has a functional dependence on the period-to-wavelength ratio. Using the basic concepts of the extended scalar theory allows for a degree of insight into the optimum design of grating structures.

A graphical comparison of the results from the three theories is useful to visualize the differences in predicting diffraction efficiencies. Figures 10 and 11 plot the predicted first-order diffraction efficiencies as a function of the wavelength-to-period ratio for substrates with refractive indices of 1.5 and 4, respectively. The grating profiles are 16 phase level approximations to the optimum continuous profiles. The gratings on the $n = 4$ substrate are assumed to have an optimum quarter-wave antireflection coating; the $n = 1.5$ substrate is uncoated.

The most important feature of Figures 10 and 11 is the significant deviation between the scalar and the other two theories for moderate wavelength-to-period ratios. The curves confirm that the scalar theory is only valid for very small wavelength-to-period ratios. Another feature illustrated in the figures is the effect of the index of refraction of the substrate. Higher-index substrates suffer a smaller diffraction efficiency falloff than do low-index substrates. This effect is readily explained from the light shadowing concept presented in Section 4.

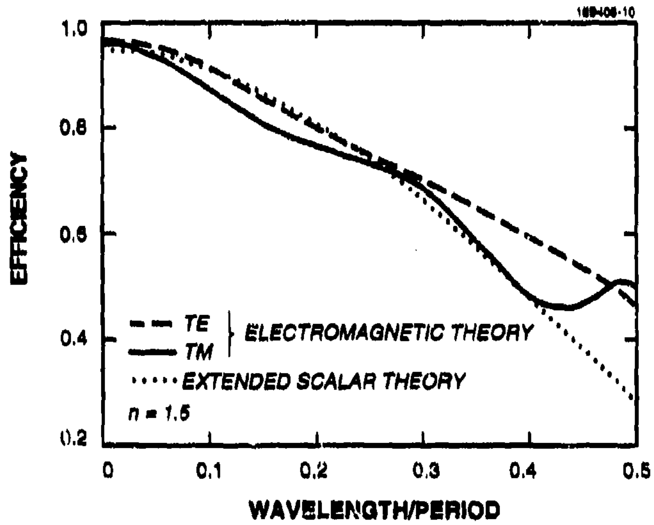


Figure 10. Predicted first-order diffraction efficiency as a function of the wavelength-to-period ratio for a grating on a substrate with $n = 1.5$.

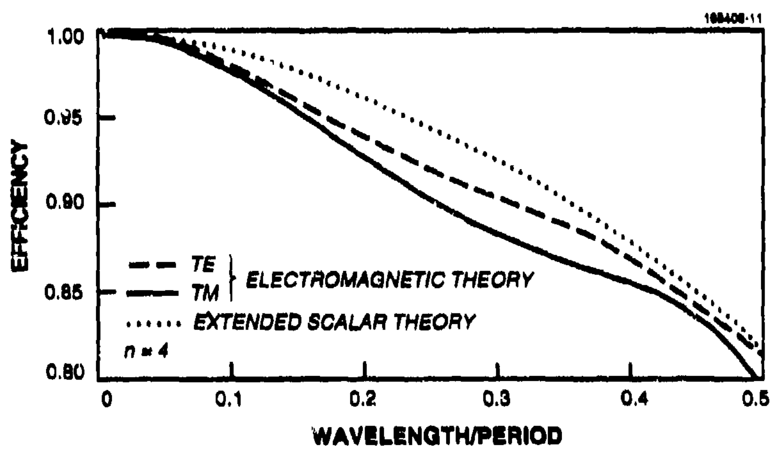


Figure 11. Predicted first-order diffraction efficiency as a function of the wavelength-to-period ratio for a grating on a substrate with $n = 4$.

It has been noted that the diffraction efficiency results of the extended scalar and the electromagnetic theories are dependent on the period-to-wavelength ratio; therefore, diffractive structures more complicated than simple periodic gratings have diffraction efficiencies that are a function of position on the element. Assigning a single diffraction efficiency value to an element requires sampling the aperture.

The diffraction efficiency of a diffractive lens, for example, can be approximately determined by assigning a periodicity to the lens that is a function of radial position. The lens can then be divided into annular regions of equal area. Each annular region is assigned a period equal to the period at its center. The extended scalar or the electromagnetic theory can then be used to determine the approximate diffraction efficiency of the annular regions. Since each region is of equal area, the lens can be assigned a diffraction efficiency that is simply the average of all the efficiencies of the annular regions. The accuracy of this approach is determined mainly by the number of annular regions into which the lens is segmented.

Using the approach described above, a first-order diffraction efficiency can be assigned to a diffractive lens as a function of its numerical aperture. Figures 12 and 13 plot the theoretical diffraction efficiencies as a function of numerical aperture for substrates with indices of refraction of 1.5 and 4, respectively. The substrate with an index of refraction of 4 is, as in the previous calculations, assumed to have an antireflection coating. The substrate with an index of refraction of 1.5 is uncoated. Curves are plotted from calculations of the electromagnetic and the extended scalar theories.

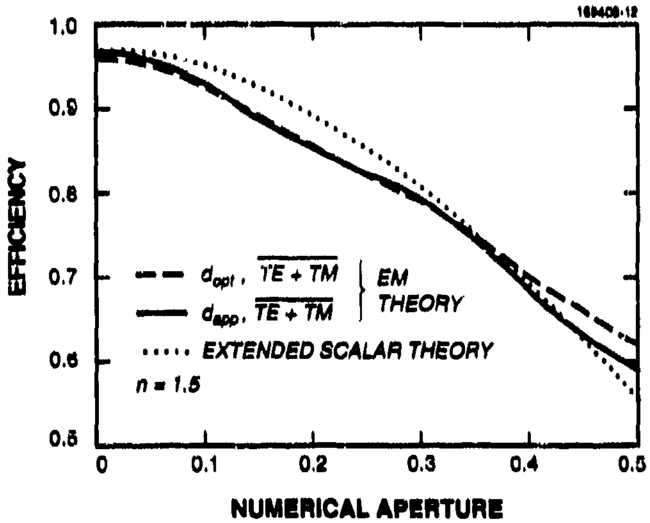


Figure 12. Predicted first-order diffraction efficiency of a diffractive lens as a function of numerical aperture for a substrate with n = 1.5.

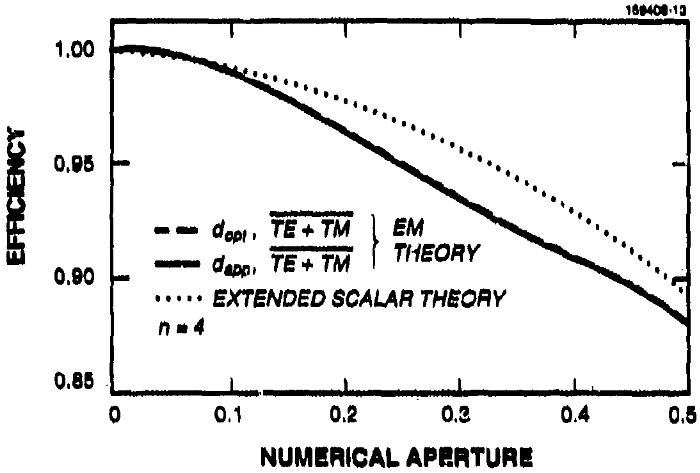


Figure 13. Predicted first-order diffraction efficiency of a diffractive lens as a function of numerical aperture for a substrate with n = 4.

The extended scalar theory calculations in Figures 12 and 13 were done for diffractive lenses with an optimum depth, while the electromagnetic theory calculations were done for diffractive lenses that had optimum depth profiles, as well as the approximate depth, determined from the scalar theory. Since optimum depth is a function of period, it varies for a lens as a function of radial position. Diffractive lenses with radially varying depths cannot realistically be fabricated using lithographic techniques; however, they can be produced using diamond turning methods.

Another difference is that the extended scalar theory is polarization independent, while the electromagnetic theory is dependent on the polarization of the incident light. On a radially symmetric diffractive lens, different angular positions are illuminated with different polarizations. The net effect over the entire aperture is simply an average of the diffraction efficiencies of the tranverse electric (TE) and transverse magnetic (TM) polarization states.

The main point elucidated in Figures 12 and 13 is that the diffraction efficiency from a diffractive lens is theoretically limited. The difference in efficiency between that predicted from the scalar theory and that predicted from a more accurate theory is dependent on the numerical aperture of the lens, and the difference becomes quite large as the numerical aperture increases. Diffraction efficiency is also a function of the index of refraction of the substrate. Diffractive lenses of a given numerical aperture have a higher theoretical efficiency on high-index substrates than on low-index substrates.

REFERENCES

1. M.W. Farn and J.W. Goodman, "Effect of VLSI fabrication errors on kinoform efficiency," *SPIE Proc.* Vol. 1211, 125–132 (1990).
2. J.A. Cox, T.R. Werner, J.C. Lee, S.A. Nelson, B.S. Fritz, and J.W. Bergstrom, "Diffraction efficiency of binary optical elements," *SPIE Proc.* Vol. 1211, 116–124 (1990).
3. J.W. Goodman, "Introduction to Fourier Optics," New York: McGraw-Hill (1968).
4. W.H. Lee, "Computer-generated holograms: Techniques and applications," in *Progress in Optics* **16**, E. Wolf (ed.), 119–232 (1978).
5. M.G. Moharam and T.K. Gaylord, "Diffraction analysis of dielectric surface-relief grating," *J. Opt. Soc. Am.* **72**, 1383–1392 (1982).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1 March 1991	Technical Report

4. TITLE AND SUBTITLE

Binary Optics Technology: Theoretical Limits on the Diffraction Efficiency of Multilevel Diffractive Optical Elements

5. FUNDING NUMBERS

C — F19628-90-C-0002
PE — 62702E
PR — 305

6. AUTHOR(S)

Gary J. Swanson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

TR-914

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ESD-TR-90-188

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Theoretical constraints limit the diffraction efficiency obtainable from multilevel diffractive optical elements. The scalar theory, commonly used to predict diffraction efficiencies, is overly optimistic. An extension to this theory is presented and compared with rigorous electromagnetic theory calculations. The extended scalar theory adds a degree of intuition in understanding why the diffraction efficiency of these elements is limited.

14. SUBJECT TERMS

binary optics
diffractive optical elements

15. NUMBER OF PAGES

38

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR